Run II Computing Review

Offsite Computing

Rick St. Denis – Glasgow University
Stefano Belforte – INFN Trieste

OutLine

- Bottom Line
- Present
- Present and Future:SAM for CDF
- Future I
- Future II
- Use Case: J/q
- Numerology
- Conclusion

Bottom lines

Run I

- CDF heavily centralized in Run1;
- Computing limited physics output.

Run II

- New Possibilities  new loads 
- Run II: No Five Year Plan more capitalistic approach. 
- The Grid cometh and must have a Fabric.
- Morph to LHC

Bottom lines

- First thought: Offload Fnal batch

  - BUT this really puts a larger load on the FNAL computing resources.
  - Leverages World-wide resources: More Physics output from detector/accelerator investment

Bottom Lines

- Treating WAN as an abundant file transfer resource.
  - Totally stupid in Run 1.
  - Expected in Run 2.

Present situation

CDF Grid

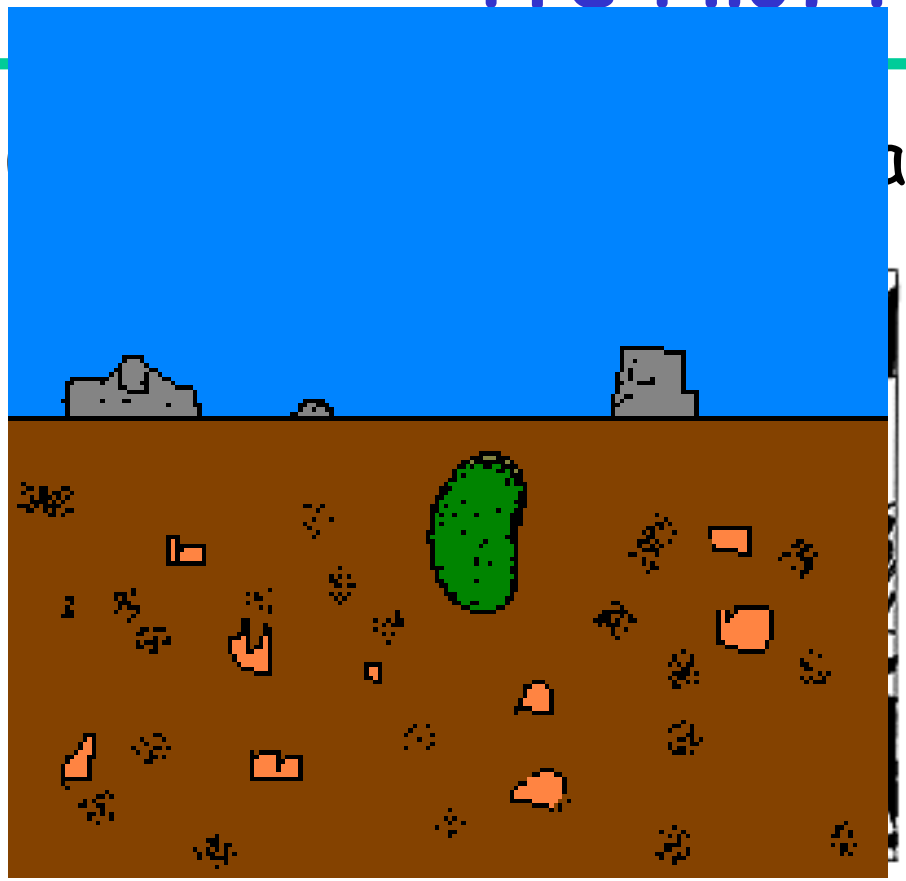
- US sites mostly planning for
  - 2ndary data copy by hand
  - Direct plug to production output
- UK and Italy:
  - 2ndary data copy via SAM
  - UK: Integration of sites via GRID O(1/10 CAF)
  - Italy: Large national central analysis farm O(1/10 CAF) accessed via GRID
- Spain/Germany: similar to UK/Italy
- Korea: exploit local large farm being built for LHC Grid

SAM for CDF

- History: UK GridPP funding: 4 seed posts for CDF and D0
  - Monte Carlo, SAM for CDF, 2 to for Gridification: build Grid into SAM while influencing Grid with SAM (Gabriele's Plans define this).
  - Risk and Resource assessment: Define 3 phases: pilot, integration, Grid-ification
- Enormous Thanks to CD, D0-CD, CDF-CD for help!

Pre Pilot Project

- [ake Oil Test]



- Adapt to CDF software - AC++ interface: Done
- Access to Enstore: Done
- Database and Server Installation: Done
- Distribution of Stations:

Difficult: SAM for CDF project

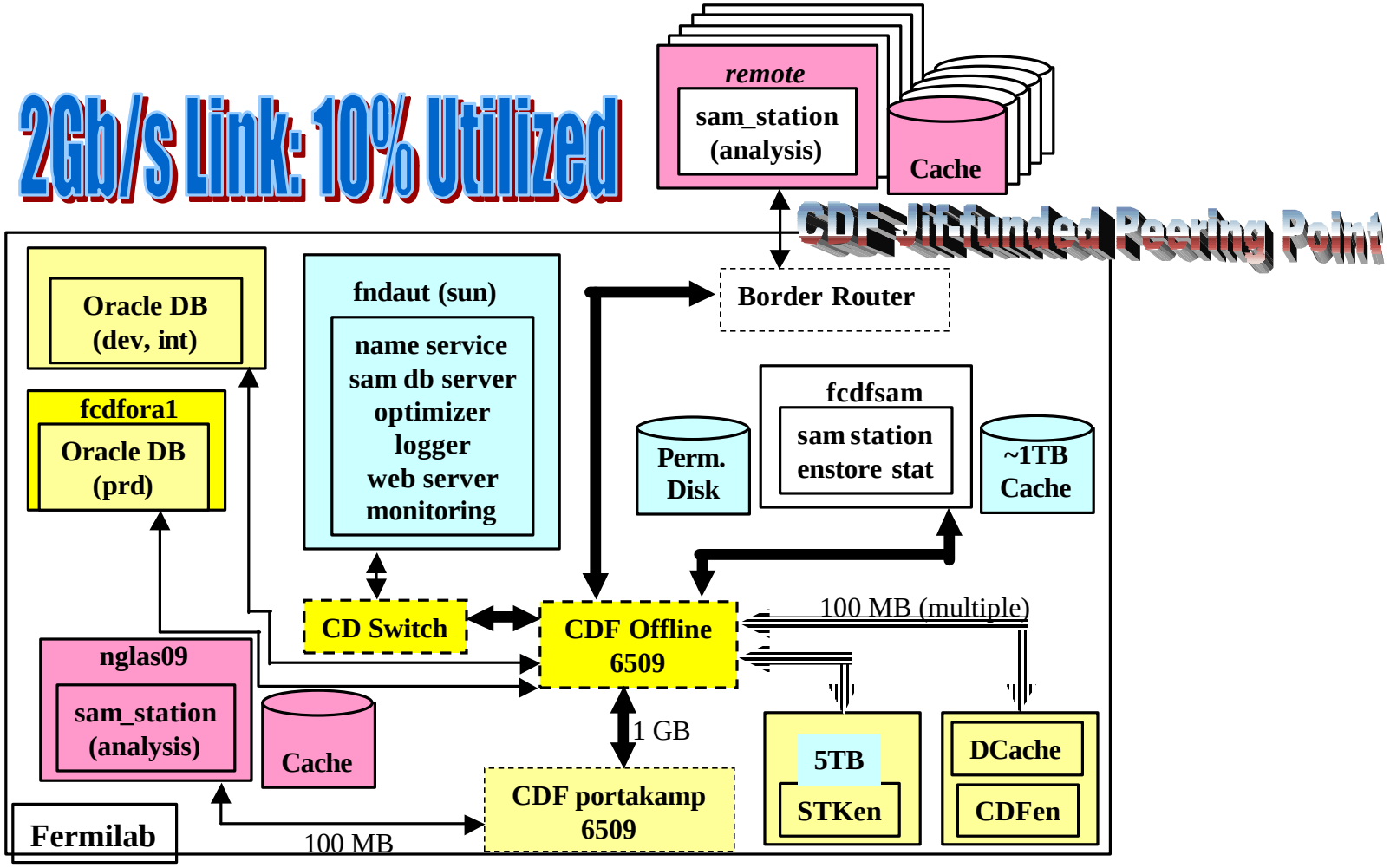
SAM for CDF

- Very strong Support of D0 and CD,
- Grew from Pre-Pilot
- June 2002
- Goals
  - Enable 5 groups to do physics analysis offsite
  - Enable access to datasets of interest
  - Production availablility of SAM (24x7)
  - Limit impact on enstore but use CDFen

Status of SAM for CDF

- Hardware and Software infrastructure in place
- DFC to SAM translation in production (Now)
- Developed AC++ interfaces to SAM to retrieve and analyze files. Automatic output to SAM not ready.
- Enabled access to DCache.
- Deploying to test sites to sort out configuration issues.
- Test users are now using SAM for physics.

Stress Existing Hardware: Upgrade



Future scenarios

- Large MC production and/or special reprocessing offsite
  - Will have some data import as well, even if unplanned
- SAM should become "the CDF data export tool" developed in a joint project with CD/D0
  - Making data export easier, will rise demand
  - We plan on SAM/Grid 

- Move data to the jobs or jobs to the data funded in collaboration of GridPP, EDG, DataTAG, PPDG.

Analysis Now: A real use case

- Now
  - J/Psi Dataset skimmed, copied to Glasgow by hand.
  - Rerun Tracking, new constants, 8way
  - No bookkeeping.
  - Need to store both input and output, doubling remote disk needs
  - Output only available at one remote site,Glasgow, and no cache.

Analysis Now: A real use case

- Test for the next now.
  - Last week "analyzed" 12 files: took all morning to transfer
  - 5s to copy if 2Gb/s! But single threading of FTP, bottlenecks.

Analysis Now: A real use case

- Next reprocess 500 G 2 tracks trigger with same exe: quality check of complex sample from Jpsi
- Then to lepton plus displaced vertex…
- Numbers in estimate are VERY volatile
  - 8 processor CPU: 6 min/file, 4K evt/file, 100 files: 10h
  - 64 dual processors if 2wk of work

Analysis Now: A real use case

- With SAM (over next couple of weeks)
  - Get each file: 50 min to reconstruct, 5s to copy?!

  - Solid bookkeeping of version of executable, calibration constants.
  - No need to store input: discard as processed: halves disk needs.
  - Store Files back at FNAL for all users to get. Files cached for efficient use.

Future scenarios II

- Sharing of resources FNAL/UK/Italy/Korea/….
  - Never a plan
  - Now a possibility
  - Politics a brand new field to explore: how to monetize ?
  - Want to make sure it feasible as soon as tools and and talks are in place: Opportunity to DEFINE a Grid for running experiments – The only way to make sure grid is useful for anyone is to make sure it is useful for someone.
  - GRIDPP: These experiments will be the source of pressure (John Gordon, RAL, May 2002, GRIDPP collab. Mtg.)

Future scenarios II

- The only thing I can be sure of:
  - I can have dedicated CPU cycles at home
    - ☞ Therefore I need to copy the data here
  - If data are too much
    - ☞ Still I need to copy output at home
    - ☞ With SAM, I can copy input to home and don't have to run at FNAL. Will run where it is best to run.
    - ☞ But that may be FNAL!
  - The GRID MUST have a fabric! It will seek the best CPU- Worst case: what I HAVE myself at home and at FNAL.

4 ways of estimating needed WAN: 1,2

- 1. Survey says: 50 remote sites each have 1.5TB storage, refresh every 2 months: 75Mbit/sec in 2002
  - doubling rate each year, in 5 years at 1 Gb/s
- 2. Every institution copies all DST as they are produced at 2.5MB/sec, x50 instit: 1Gbit/sec flat
- 3. Grid scenario: 1.6Gbit/sec
  - Data sets copied offsite Eg. Via SAM.
  - rCAF users analyze at 10 MB/sec
  - 20% of local cache miss. Get 2MByte/sec/user.
  - 100 offsite users: is 1.6 Gbit/sec.

4 ways of estimating needed WAN: 4

- 4: Produce ntuple at fnal and copy home.
  - Don't copy O(10TB) home, leave at Fnal, run on Fnal-CAF
  - Let Ntuple be 1/100 of input: 1TB Run2 dataset produces 10GB for home desktop (= ~cdfsga)
  - Jobs run at few MB/sec (CAF Spec)
  - every user is expected to run O(10) jobs in parallel
  - Let's assume 100Mbyte/sec/user of input.
    - Output is thus 1MByte/sec/user
    - x200 active users = 1.6 Gbit/sec
  - This is for Run2a sizes. As data grow will have faster CPU, faster I/O and/or more parallelism.

WAN summary

- 4 ways of counting: Each says 2Gbit/sec
- We trust the latter 2: Likely both, but different people
- Can imagine 3 Gbit/sec by 2004 (Run2a) but most of the estimates scale with people:
  - May need 3 Gbit/sec sooner if development is over and can turn to analysis
- WAN has to be 30% occupied at most to work well
- FNAL need O(10Gbit/sec) only for CDF
- Those numbers hold for 2004
- Then double every year (for last 2 est.)

Conclusion

- Run II: Brave New World. More efficient Physics Yield with outside resources: stress FNAL batch, networking and support resource needs: Boomerang effect: The Grid needs fabric!
- GRIDPP admits the experiments will pressure the GRID: running experiments (BaBar, CDF, D0) to be featured in September 2002 collaboration meeting.
- CDF-EuroGrid Workshop in September.(add US, D0)
- LHC pressure: better be doing same thing.
- The Grid will happen. Particle physics can define the grid. FNAL can define the grid.

To work for Anyone it must work for Someone!

- If there is a fork in the road, take it.